

## University of Groningen

### **Large normal-range TBP and ATXN7 CAG repeat lengths are associated with increased lifetime risk of depression**

Gardiner, S. L.; van Belzen, M. J.; Boogaard, M. W.; van Roon-Mom, W. M. C.; Rozing, P.M.; van Hemert, A. M.; Smit, J. H.; Beekman, A. T. F.; van Grootheest, Gerard; Schoevers, R. A.

*Published in:*  
Translational Psychiatry

*DOI:*  
[10.1038/tp.2017.116](https://doi.org/10.1038/tp.2017.116)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2017

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Gardiner, S. L., van Belzen, M. J., Boogaard, M. W., van Roon-Mom, W. M. C., Rozing, P. M., van Hemert, A. M., Smit, J. H., Beekman, A. T. F., van Grootheest, G., Schoevers, R. A., Voshaar, R. C. O., Comijs, H. C., Penninx, B. W. J. H., van der Mast, R. C., Roos, R. A. C., & Aziz, N. A. (2017). Large normal-range TBP and ATXN7 CAG repeat lengths are associated with increased lifetime risk of depression. *Translational Psychiatry*, 7, [e1143]. <https://doi.org/10.1038/tp.2017.116>

#### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

#### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

## ORIGINAL ARTICLE

Large normal-range *TBP* and *ATXN7* CAG repeat lengths are associated with increased lifetime risk of depression

SL Gardiner<sup>1,2</sup>, MJ van Belzen<sup>3</sup>, MW Boogaard<sup>1,3</sup>, WMC van Roon-Mom<sup>2</sup>, MP Rozing<sup>4</sup>, AM van Hemert<sup>5</sup>, JH Smit<sup>6</sup>, ATF Beekman<sup>6</sup>, G van Grootheest<sup>6</sup>, RA Schoevers<sup>7</sup>, RC Oude Voshaar<sup>7</sup>, HC Comijs<sup>6</sup>, BWJH Penninx<sup>6</sup>, RC van der Mast<sup>5,8</sup>, RAC Roos<sup>1</sup> and NA Aziz<sup>1</sup>

Depression is one of the most prevalent and debilitating psychiatric disorders worldwide. Recently, we showed that both relatively short and relatively long cytosine–adenine–guanine (CAG) repeats in the huntingtin gene (*HTT*) are associated with an increased risk of lifetime depression. However, to what extent the variations in CAG repeat length in the other eight polyglutamine disease-associated genes (PDAGs) are associated with depression is still unknown. We determined the CAG repeat sizes of *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *TBP*, *ATN1* and *AR* in two well-characterized Dutch cohorts—the Netherlands Study of Depression and Anxiety and the Netherlands Study of Depression in Older Persons—including 2165 depressed and 1058 non-depressed individuals—aged 18–93 years. The association between PDAG CAG repeat size and the risk for depression was assessed via binary logistic regression. We found that the odds ratio (OR) for lifetime depression was significantly higher for individuals with > 10, compared with subjects with ≤ 10, CAG repeats in both *ATXN7* alleles (OR = 1.90, confidence interval (CI) 1.26–2.85). For *TBP* we found a similar association: A CAG repeat length exceeding the median in both alleles was associated with an increased risk for lifetime depression (OR = 1.33, CI 1.00–1.76). In conclusion, we observed that carriers of either *ATXN7* or *TBP* alleles with relatively large CAG repeat sizes in both alleles had a substantially increased risk of lifetime depression. Our findings provide critical evidence for the notion that repeat polymorphisms can act as complex genetic modifiers of depression.

*Translational Psychiatry* (2017) 7, e1143; doi:10.1038/tp.2017.116; published online 6 June 2017

## INTRODUCTION

Depression is one of the most common psychiatric disorders with an estimated lifetime prevalence of 14.6% in high-income and 11.1% in low-income countries.<sup>1</sup> Worldwide, depression is responsible for 74.5 million disability-adjusted life years<sup>2</sup> and in 2010 the total economic burden of depression in the United States alone was estimated at \$210.5 billion,<sup>3</sup> indicating that depression is a major strain on society. In order to create new and more effective therapeutic and preventive strategies, unravelling the pathogenesis of depression is imperative and elucidation of its genetic determinants is a critical step in this process.

The heritability of major depressive disorder (MDD) has been estimated to be between 30 and 50%.<sup>4</sup> Nevertheless, genome-wide association studies (GWASs) have shown limited success in identifying its genetic basis. In a mega-analysis with more than 18 000 subjects from European ancestry, no single-nucleotide polymorphism (SNP) reached genome-wide significance.<sup>4</sup> However, recently, a meta-analysis of results from three studies, together containing 180 866 individuals, found two lead SNPs associated with depression.<sup>5</sup> In addition, another study combining three cohorts in a joint analysis with a total of 478 240 subjects found 17 independent SNPs significantly associated with a diagnosis of MDD,<sup>6</sup> suggesting that GWASs can successfully identify genetic associations with highly polygenic phenotypes.

However, the identified effect sizes were small. Moreover, to gain sufficient statistical power the sample sizes required were extremely large and, as a consequence, unavoidably heterogeneous. Therefore, although GWASs have contributed greatly to genetic mapping of complex human traits, their success has been limited by the fact that, aside from SNPs, these studies are unsuitable to assess the contribution of other genetic polymorphisms, especially DNA repeat sequences.<sup>7</sup>

Tandem repeats constitute ~3% of the human genome, a higher percentage than the entirety of the protein-coding sequences,<sup>8</sup> substantially contributing to genetic variation.<sup>9,10</sup> Many rare hereditary disorders are caused by expansions of simple DNA repeat sequences.<sup>11</sup> However, the association of DNA repeat sequences with more common diseases is largely unknown. Polyglutamine diseases are the most prevalent disorders caused by an expanded DNA repeat sequence.<sup>11,12</sup> These diseases are caused by an expansion of a trinucleotide repeat (cytosine–adenine–guanine (CAG)) in the translated region of various genes. As CAG encodes the amino acid glutamine, the expansion in the trinucleotide repeat sequence results in an elongated polyglutamine domain in the associated proteins.<sup>12</sup> The most common polyglutamine disease is Huntington disease, a severe neurodegenerative disorder characterized by both motor and neuro-psychiatric impairment. Huntington disease is caused by a CAG

<sup>1</sup>Department of Neurology, Leiden University Medical Centre, Leiden, The Netherlands; <sup>2</sup>Department of Human Genetics, Leiden University Medical Centre, Leiden, The Netherlands; <sup>3</sup>Department of Clinical Genetics, Leiden University Medical Centre, Leiden, The Netherlands; <sup>4</sup>Centre for Healthy Ageing/Department of Public Health, Section of Social Medicine, University of Copenhagen, Copenhagen, Denmark; <sup>5</sup>Department of Psychiatry, Leiden University Medical Centre, Leiden, The Netherlands; <sup>6</sup>Department of Psychiatry, Amsterdam Public Health Research Institute and Neuroscience Campus Amsterdam, VU University Medical Centre/GGZ inGeest, Amsterdam, The Netherlands; <sup>7</sup>Department of Psychiatry, University of Groningen, University Medical Centre Groningen, Research School Cognitive Behavioural Neuroscience, Groningen, The Netherlands and <sup>8</sup>Department of Psychiatry, Collaborative Antwerp Psychiatric Research Institute (CAPRI), University of Antwerp, Antwerp, Belgium. Correspondence: SL Gardiner, Department of Neurology, Leiden University Medical Centre, Einthovenweg 20, PO Box 9600, Postzone S-4-P, Leiden 2300 RC, The Netherlands. E-mail: s.l.gardiner@lumc.nl

Received 11 April 2017; accepted 20 April 2017

**Table 1.** Summary genotyped polyglutamine disease-associated genes

Gene	Disease	Protein	CAG repeat ranges		Allele	Mean	Median	N	Range	No. of intermediate <sup>a</sup>	No. of pathological <sup>a</sup>
			Normal	Pathological							
ATXN1	SCA1	Ataxin-1	6–39	41–83	Short	29.22	29.00	3228	18–35	0	0
					Long	30.76	30.00	3228	26–39	0	0
ATXN2	SCA2	Ataxin-2	14–32	33–500	Short	21.91	22.00	3236	14–24	0	0
					Long	22.35	22.00	3236	22–36	0	2 (33, 36)
ATXN3	SCA3	Ataxin-3	12–44	52–87	Short	18.96	20.00	3235	13–33	0	0
					Long	24.27	23.00	3235	13–49	1 (49)	0
CACNA1A	SCA6	CACNA1A	4–18	20–33	Short	10.52	11.00	3220	4–14	0	0
					Long	12.45	13.00	3220	4–22	0	1 (22)
ATXN7	SCA7	Ataxin-7	3–19	37–460	Short	10.04	10.00	2979	7–15	0	0
					Long	10.85	10.00	2979	7–19	0	0
TBP	SCA17	TBP	25–43	45–66	Short	36.31	37.00	3238	27–39	0	0
					Long	37.89	38.00	3238	32–48	1 (44)	1 (48)
ATN1	DRPLA	Atrophin-1	3–38	48–93	Short	12.36	14.00	3228	5–22	0	0
					Long	15.47	15.00	3228	8–30	0	0
AR	SBMA	Androgen receptor	6–36	38–72	Short	20.80	21.00	3228	8–36	0	0
					Long	22.90	23.00	3228	8–36	0	0

Abbreviations: CACNA1A, calcium channel, voltage-dependent P/Q type,  $\alpha$  1A subunit; CAG, cytosine–adenine–guanine; DRPLA, dentatorubropallidoluysian atrophy; SBMA, spinal bulbar muscular atrophy; SCA, spinocerebellar ataxia; TBP, thymine–adenine–thymine–adenine (TATA) box-binding protein.

repeat expansion in exon 1 of the huntingtin (*HTT*) gene.<sup>13,14</sup> Recently, we demonstrated that both relatively short and relatively large CAG repeats in the longer *HTT* allele are associated with lifetime depression, suggesting that repeat polymorphisms could also act as complex genetic modifiers of depression and account for part of its ‘missing heritability’.<sup>15</sup>

Apart from Huntington disease, eight other polyglutamine disorders, all of which are neurodegenerative disorders, are frequently associated with considerable neuropsychiatric impairment (Table 1).<sup>16–35</sup> These rare disorders result from relatively large repeat expansions in polyglutamine disease-associated genes (PDAGs). However, to what extent more common repeat length variations in the normal range of these genes could act as genetic modifiers of depression in the general population is still unknown. Hence, here we aimed to assess the contribution of CAG repeat length variations in these other PDAGs to depression susceptibility using data from two well-defined Dutch cohorts: The Netherlands Study of Depression and Anxiety (NESDA) and the Netherlands Study of Depression in Old Persons (NESDO).

## MATERIALS AND METHODS

We genotyped eight PDAGs (*ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATXN7*, *TBP*, *ATN1* and *AR*) in all participants with sufficient amounts of DNA available from blood samples of two well-characterized Dutch cohorts: the NESDA (cohort 1) and the NESDO (cohort 2) cohorts (Table 1 and Supplementary Figures 1–17).

### Cohort 1

The NESDA is a cohort study among 2981 participants aged 18–65 years.<sup>36</sup> The participants were recruited from the general population, general practices and mental health-care institutes. The sample included 1973 subjects with a lifetime diagnosis of depression and/or dysthymia (including 1925 patients with MDD), 635 subjects with a lifetime anxiety disorder without lifetime depression and 373 healthy controls.<sup>36</sup> Diagnoses were made in accordance with the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition criteria using the WHO Composite International Diagnostic Interview.<sup>37</sup>

### Cohort 2

The NESDO is a cohort study among 510 participants aged 60–93 years.<sup>38</sup> The participants were recruited from both general practices and mental

health-care institutes. The sample included 378 depressed individuals (360 with MDD) and 132 healthy controls. The same methods for diagnosing depression were used as in the NESDA.<sup>38</sup>

### Genotyping

A polymerase chain reaction (PCR) was performed in a TProfessional thermocycler (Biometra, Westburg, Leusden, the Netherlands) with labelled primers flanking the CAG stretch of the PDAGs (Biolegio, Nijmegen, the Netherlands; Supplementary Table 1). The PCR was performed using 10 ng of genomic DNA, 1 × OneTaq mastermix (New England Biolabs, Ipswich, MA, USA, OneTaq Hot start with GC Buffer mastermix), 1 µl of primer Mix A or B (Supplementary Table 1) and Aqua B. Braun water to a final volume of 10 µl. The PCR was run with 27 cycles of 30 s, denaturation at 94 °C, 1 min of annealing at 60 °C and 2 min elongation at 68 °C, preceded by 5 min of initial denaturation at 94 °C. Final elongation was performed at 69 °C for 5 min. Every PCR included a negative control without genomic DNA and a reference sample of CEPH 1347-02 genomic DNA. The PCR products were run on an ABI 3730 automatic DNA sequencer (Applied Biosystems, Foster City, CA, USA) and analysed using the GeneMarker software version 2.4.0. (SoftGenetics, State College, PA, USA). For every analysis, we included three controls with known CAG repeat lengths for each PDAG to assure that every run was performed reliably. All assessments were made with cases and controls randomized on plates and blinding with respect to disease status information.

### Statistical analysis

Binary logistic regression was used to assess whether CAG repeat sizes in the two alleles of each PDAG were associated with the risk of lifetime depression. For each PDAG, in an initial model, the presence of lifetime depression (that is, MDD and/or dysthymia) was set as the dependent variable and the CAG repeat lengths of both alleles were used as the independent variables. To assess interaction effects between the two alleles or nonlinear effects, a product term of the two alleles and a quadratic term for each allele were added to the model. In case the product term or the quadratic terms were not significant, these predictors were removed from the model and the analysis was repeated. Subsequently, we adjusted the results for the effects of sex, age and education level (coded as ‘basic’, ‘intermediate’ and ‘high’<sup>36,38</sup>) in order to assess whether the effect of CAG repeat size variations was independent of these well-established risk factors for depression. The Nagelkerke  $R^2$  was used to assess the proportion of variability explained by the predictors in the model. To account for potential effects of heteroscedasticity and influential points, all statistical significance tests were based on robust estimators of s.e.’s. Moreover, to visualise our results as well as to assure

**Table 2.** The distribution of participants in three ATXN7 categories

ATXN7 category	Allele	Mean	Range	Lifetime depression <sup>a</sup>	No lifetime depression <sup>b</sup>	Odds ratio (95% CI)
I <sup>c</sup>	Short	9.94	7–10	1220	619	1.00
	Long	10.00	7–10			
II	Short	9.97	7–10	662	331	1.01 (0.86–1.20)
	Long	12.18	11–19			
III	Short	11.90	11–15	116	31	(1.26–2.85) <sup>d</sup>
	Long	12.54	11–18			

Abbreviations: I, both alleles of ATXN7 contain a CAG repeat number  $\leq 10$ ; II, the relatively longer ATXN7 allele contains a CAG repeat number  $> 10$  and the relatively shorter allele contains a CAG repeat number  $\leq 10$ ; III, both alleles of ATXN7 contain a CAG repeat number  $> 10$ ; CAG, cytosine–adenine–guanine; CI, confidence interval. <sup>a</sup>Figures represent the number of subjects per ATXN7 category in the group with lifetime depression. <sup>b</sup>Figures represent the number of subjects per ATXN7 category in the group with no lifetime depression. <sup>c</sup>This category contains the largest number of subjects and was therefore chosen as the reference category for odds ratio calculation. <sup>d</sup> $P < 0.01$  by Fisher's exact test in comparison with the reference category.

that the results were not unduly affected by violated model assumptions, we also applied a non-parametric method: we calculated the odds ratios (ORs) for a lifetime diagnosis of depression per CAG repeat group. The CAG repeat groups were defined based on the median CAG repeat size of each allele. A division as such resulted in three adequately sized groups: I, both alleles  $\leq$  their median; II, the relatively shorter allele  $\leq$  the median and the relatively longer allele  $>$  the median; III, both alleles  $>$  their median. Subsequently, the ORs were compared with the Fisher's exact test. In order to account for multiple testing, we applied a false discovery rate correction as described by Benjamini and Hochberg, assuming eight independent tests with a two-sided  $\alpha$  of 0.05.<sup>39</sup> All data are displayed as means and 95% confidence intervals (CIs), unless otherwise specified. All analyses were performed in SPSS version 23.0 (IBM SPSS Statistics for Windows, IBM, Armonk, NY, USA).

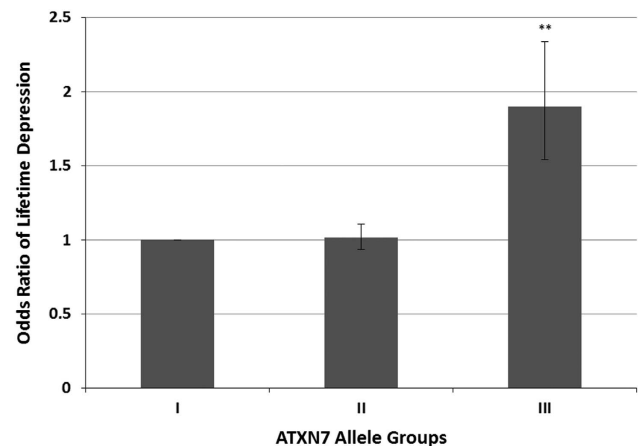
## RESULTS

Risk of lifetime depression increases with higher ATXN7 CAG repeat sizes

In total, we were able to genotype a total of 2979 participants for ATXN7 (Table 1). This included 1998 depressed subjects and 981 non-depressed subjects. For 512 individuals, too little DNA material was available to determine the CAG repeat length in ATXN7 (Supplementary Table 2). These lacking samples were missing completely at random. The CAG repeat lengths ranged from 7 to 19 repeats (Table 1).

We found a significant association between the risk of lifetime depression and the CAG repeat length of ATXN7 in the shorter allele ( $\beta = 0.202$ ,  $P = 0.006$ ). Adjusting for the effects of gender, age and education hardly changed these results ( $\beta = 0.184$ ,  $P = 0.013$ ; Supplementary Table 3). After adjusting for multiple testing, this association remained significant. Inclusion of the CAG repeat length of the ATXN7 allele in the model increased the proportion of explained genetic variation from the baseline model including only gender, age and level of education by 0.004 (that is, the  $R^2$  increased from 0.024 to 0.028). Considering our control group included subjects with anxiety disorders without comorbid depression ( $n = 348$ ), we also performed a sensitivity analysis by excluding these individuals from the control group. Neither the parameter estimates nor their significance were materially altered by this procedure ( $\beta = 0.215$ ,  $P = 0.009$ ).

For the non-parametric method, we divided all subjects based on the median CAG repeat size for each ATXN7 allele (median relatively short allele = 10, median relatively long allele = 10) and compared the ORs of lifetime depression among the groups (Table 2). The group containing the largest number of individuals was used as the reference category. Comparing the other groups to the reference category demonstrated the significant effect of the shorter allele on the risk of lifetime depression (Figure 1). When the relatively short allele and thus also the relatively long allele both contained a CAG repeat length larger than 10 repeats,



**Figure 1.** Odds ratio for lifetime depression per ATXN7 category. The odds ratio for lifetime depression increases significantly when both the relatively long ATXN7 allele and the relatively short ATXN7 allele exceed the median CAG repeat number of 10. The odds of having lifetime depression almost doubles. The group having both alleles with a CAG repeat number  $\leq 10$  was the largest and, therefore, was set as the reference category. I, both alleles of ATXN7 contain a CAG repeat number  $\leq 10$ ; II, the relatively longer ATXN7 allele contains a CAG repeat number  $> 10$  and the relatively shorter allele contains a CAG repeat number  $\leq 10$ ; III, both alleles of ATXN7 contain a CAG repeat number  $> 10$ . Error bars indicate  $\pm$  s.e.  $**P < 0.01$  by the Fisher's exact test in comparison with the reference category; CAG, cytosine–adenine–guanine.

the odds for lifetime depression almost doubled compared with the reference category in which both ATXN7 alleles had a CAG repeat number equal to or smaller than the median.

TBP CAG repeat sizes in both alleles interact to affect the risk of lifetime depression

A total of 3238 individuals were genotyped for TBP (Table 1). This included 2180 people diagnosed with lifetime depression and 1058 non-depressed people. For 253 individuals, too little DNA material was available to determine the CAG repeat length in TBP (Supplementary Table 2). The lacking samples were missing completely at random. For TBP, we found CAG repeat lengths ranging between 27 and 48 repeats (Table 1).

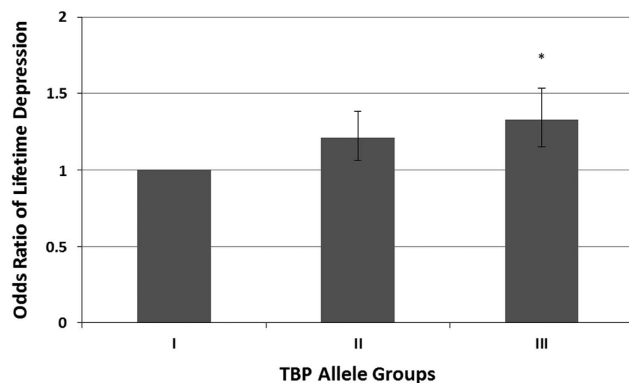
We found the interaction between the number of CAG repeats in the two TBP alleles to be significantly associated with the risk of lifetime depression (TBP short allele:  $\beta = -2.270$ ,  $P = 0.004$ ; TBP long allele:  $\beta = -2.112$ ,  $P = 0.006$ ; TBP interaction term:  $\beta = 0.060$ ,  $P = 0.005$ ). When adjusting for the effects of gender, age and



**Table 3.** The distribution of participants in three *TBP* categories

<i>TBP</i> category	Allele	Mean	Range	Lifetime depression <sup>a</sup>	No lifetime depression <sup>b</sup>	Odds ratio (95% CI)
I <sup>c</sup>	Short	35.78	27–37	1462	716	1.00
	Long	37.47	32–38			
II	Short	36.04	30–37	230	93	1.21 (0.94–1.57)
	Long	39.30	39–48			
III	Short	38.09	38–39	201	74	1.33 (1.00–1.76) <sup>d</sup>
	Long	39.28	39–44			

I, both *TBP* alleles contain a CAG repeat number  $\leq$  their median; II, the relatively longer *TBP* allele contains a CAG repeat number  $>$  the median and the relatively shorter *TBP* allele contains a CAG repeat number  $\leq$  the median; III, both *TBP* alleles contain a CAG repeat number  $>$  their median (median short allele = 37, median long allele = 38); CAG, cytosine–adenine–guanine; CI, confidence interval. <sup>a</sup>Figures represent the number of subjects per *TBP* category in the group with lifetime depression. <sup>b</sup>Figures represent the number of subjects per *TBP* category in the group with no lifetime depression. <sup>c</sup>This category contains the largest number of subjects and was therefore chosen as the reference category for odds ratio calculation. <sup>d</sup> $P < 0.05$  by Fisher's exact test in comparison with the reference category.



**Figure 2.** Odds ratio for lifetime depression per *TBP* category. The odds ratio for lifetime depression increases significantly when both the relatively shorter *TBP* allele and the relatively longer *TBP* allele exceed their median CAG repeat number (median short allele = 37, median long allele = 38). The group with both alleles being equal to or smaller than their medians was the largest and, therefore, is defined as the reference category. I, both *TBP* alleles contain a CAG repeat number  $\leq$  their median; II, the relatively longer *TBP* allele contains a CAG repeat number  $>$  the median and the relatively shorter *TBP* allele contains a CAG repeat number  $\leq$  the median; III, both *TBP* alleles contain a CAG repeat number  $>$  their median. Error bars indicate  $\pm$  one s.e. \* $P < 0.05$  by the Fisher's exact test in comparison with the reference category. CAG, cytosine–adenine–guanine; TBP, thymine–adenine–thymine–adenine (TATA) box-binding protein.

education, the results hardly changed (*TBP* short allele:  $\beta = -2.302$ ,  $P = 0.004$ ; *TBP* long allele:  $\beta = -2.150$ ,  $P = 0.005$ ; *TBP* interaction term:  $\beta = 0.061$ ,  $P = 0.005$ ; Supplementary Table 3). Furthermore, the results remained significant after performing the sensitivity analysis and applying the false discovery rate correction.<sup>39</sup> The proportion of explained genetic variation increased by 0.005 from the baseline model including only gender, age and level of education (that is, the  $R^2$  increased from 0.023 to 0.028).

For the non-parametric method, we divided all subjects in three groups based on the median CAG repeat size for each *TBP* allele (median relatively short *TBP* allele = 37, median relatively long *TBP* allele = 38) and compared the ORs of lifetime depression among the groups (Table 3). The largest group was set as the reference category. The comparison of the groups demonstrated the significant effect of the interaction between the relatively short *TBP* allele and the relatively long *TBP* allele on the risk of depression. The risk of lifetime depression was significantly higher when both *TBP* alleles had a CAG repeat length exceeding their median compared with the reference category, in which both

alleles had a CAG repeat number equal to or lower than their median (Figure 2).

CAG repeat sizes in other PDAGs were not associated with risk of lifetime depression

The number of CAG repeats in *ATXN1*, *ATXN2*, *ATXN3*, *CACNA1A*, *ATN1* and *AR* were not associated with the risk of lifetime depression. Neither the main effect of the alleles, nor the interactions between the two alleles, or the quadratic terms, were significantly associated with the presence of lifetime depression in the combined cohort. As *AR* is located on the X-chromosome, we performed one analysis stratifying the data by gender and another analysis using only either the relatively short or the relatively long allele in the model. None of the two approaches demonstrated a significant association between CAG repeat number in *AR* and depression (Supplementary Table 3).

The effects of *HTT*, *ATXN7* and *TBP* CAG repeat sizes are independent

In previous research, we demonstrated that both the main term of the longer *HTT* allele and the quadratic term of the longer *HTT* allele are also significantly associated with a diagnosis of lifetime depression.<sup>15</sup> To assess the contribution of the CAG repeat polymorphisms in all PDAGs to depression risk, we examined the degree to which *HTT*, *ATXN7* and *TBP* CAG repeat sizes explain depression heritability. To this end, we applied a multivariate model with the CAG repeat sizes in these three PDAGs as predictors. In addition, we adjusted for the effects of gender, age and education. We found that the parameter estimates as well as their associated statistical significances hardly changed, indicating that the effects of *HTT*, *ATXN7* and *TBP* CAG repeat sizes on depression susceptibility are mutually independent (*HTT* long allele:  $\beta = -0.286$ ,  $P = 0.018$ ; *HTT* long allele quadratic term  $\beta = 0.006$ ,  $P = 0.017$ ; *ATXN7* short allele:  $\beta = 0.210$ ,  $P = 0.005$ ; *TBP* short allele:  $\beta = -2.728$ ,  $P = 0.001$ ; *TBP* long allele,  $\beta = -2.566$ ;  $P = 0.001$ ; *TBP* interaction term:  $\beta = 0.072$ ,  $P = 0.001$ ). Inclusion of the CAG repeat sizes in *HTT*, *ATXN7* and *TBP* in the model increased  $R^2$  by 0.014 from the baseline model including only gender, age and level of education (from 0.024 to 0.038), indicating that the CAG repeat lengths in these PDAGs can account for an additional 1.4% of the genetic variation on the observed probability scale. We also derived the  $R^2$  on the liability scale as described previously.<sup>40</sup> Assuming that depression has a lifetime prevalence of  $\sim 15\%$  in the Netherlands<sup>41</sup> and adjusting for the oversampling of patients with depression in our cohort, the  $R^2$  on the liability scale was 0.0191, indicating that CAG repeat size polymorphisms in *HTT*, *ATXN7* and *TBP* together can account for  $\sim 1.9\%$  of depression heritability.

The prevalence of intermediate and pathological PGDA alleles. In total, four genotyped subjects had a CAG repeat number in the pathological range of a PDAG. One individual had a CAG repeat number of 22 in *CACNA1A*, one subject had a CAG repeat number of 48 in *TBP* and two people had a CAG repeat number of 33 and 36 in *ATXN2* (Table 1). All four were depressed, but at the ages of 53 years for *CACNA1A*, 59 years for *TBP*, and 26 and 31 years for *ATXN2*, none had been diagnosed with the respective diseases (that is, SCA6, SCA17 and SCA2).

One person had a CAG repeat number in the longer *TBP* allele of 44, belonging to the intermediate range. This individual was also diagnosed with lifetime depression. Furthermore, one depressed individual had an intermediate CAG repeat number in the longer allele of *ATXN3* (that is, 49 repeats). For *ATXN1*, *ATXN7*, *ATN1* and *AR* all subjects had a CAG repeat number within the normal range (Table 1).

When we included the two depressed individuals with a CAG repeat length in the reduced penetrance range of *HTT* found previously,<sup>15</sup> we observed a trend for a higher proportion of subjects with a CAG repeat in the intermediate or pathological range of a PDAG in the depressed group ( $P=0.059$ ).

## DISCUSSION

To our knowledge, this is the first study assessing the influence of CAG repeat size variations in the normal range of PDAGs—other than *HTT*—on the risk of lifetime depression. Interestingly, we found an association between the presence of lifetime depression and the CAG repeat length in two PDAGs, that is, *ATXN7* and *TBP*. The main relationship we found was that when the CAG repeat number of both alleles in either *ATXN7* or *TBP* was relatively large, the odds for a diagnosis of lifetime depression markedly increased. Moreover, six genotyped individuals had a CAG repeat number in at least one of their PDAGs that extended into the intermediate or pathological range (two intermediate, four pathological). All of these subjects were diagnosed with lifetime depression. These findings are in support of the hypothesis that repeat polymorphisms may act as complex genetic modifiers of depression and thus could account for part of its 'missing heritability'.

In our study, we found that the odds of lifetime depression almost doubled in individuals with a CAG repeat length of  $>10$  in both *ATXN7* alleles. Similarly, we found that the risk of lifetime depression markedly increased ( $OR=1.33$ ) in individuals who had a CAG repeat length exceeding the median in both *TBP* alleles. These effect sizes are considerable, especially when compared with the effect sizes of the two most significant SNPs found in a recent GWAS:  $OR=0.955$  (95% CI 0.943–0.968) and  $OR=1.051$  (95% CI 1.036–1.067).<sup>6</sup> Although we cannot fully exclude potential modifying effects of SNPs in linkage disequilibrium (LD) with *ATXN7* and *TBP*, the fact that such SNPs in these PDAGs have not been detected before in GWASs suggests that the influence of other genetic variants in LD with the CAG repeat size in these genes is likely to be minimal.<sup>42</sup> Furthermore, we found that CAG repeat size polymorphisms in *HTT*, *ATXN7* and *TBP* together can account for  $\sim 1.9\%$  of depression heritability. Using the same population prevalence of 15% for depression, the meta-analysis in the recent GWAS calculated a heritability of 5.9%.<sup>6</sup> The amount of variance explained by these SNPs is larger than the amount explained by the CAG repeat polymorphisms we found. However, the meta-analysis in this GWAS examined SNPs in the entire genome of 326 113 individuals,<sup>6</sup> whereas our estimation was based on the effects of CAG repeat variations in only nine PDAGs. This fact suggests that investigating additional repeat polymorphisms in the human genome could lead to the identification of many more novel genetic determinants of depression. The importance of investigating repeat polymorphisms in association

with health and disease has been described in previous literature and begins to gain more recognition within the field of genetic research.<sup>9,10,43–46</sup>

CAG repeat numbers exceeding 36 in *ATXN7* are responsible for the severe neurodegenerative disorder spinocerebellar ataxia type 7 (SCA7).<sup>34</sup> SCA7 is a progressive autosomal-dominant neurodegenerative disorder primarily characterized by cerebellar ataxia and macular degeneration.<sup>47</sup> Although depressive symptoms are a frequent finding in many neurodegenerative disorders, the prevalence of depression has hardly been assessed in SCA7.<sup>48–51</sup> Our findings suggest that depression might be an underappreciated feature of SCA7, which needs further characterisation. A CAG repeat number in *TBP* larger than 48 causes the severe progressive neurodegenerative disorder SCA17.<sup>34</sup> Aside from cerebellar ataxia, dementia and pyramidal symptoms, in 67% of the cases SCA17 is also accompanied by psychiatric signs and symptoms, including depression, behavioural changes as well as psychosis.<sup>52</sup> Initial symptoms of SCA17 have been described to include depression.<sup>53</sup>

The gene *ATXN7* encodes the protein ataxin-7. Apart from several important cellular functions,<sup>51</sup> ataxin-7 is an integral part of the TATA-binding protein-free TAF-containing complex.<sup>54</sup> TATA-binding protein-free TAF-containing complex allows for the initiation of transcription via RNA polymerase II in the absence of the RNA polymerase II transcription factor D.<sup>55</sup> Interestingly, the TATA-box-binding protein (TBP) that is encoded by *TBP*, is the DNA-binding subunit of RNA polymerase II transcription factor D. TBP anchors RNA polymerase II transcription factor D to the TATA-box upstream of the first codon, allowing for the initiation of transcription via RNA polymerase II.<sup>56</sup> Thus, TBP as well as ataxin-7 seem to have a part in the initiation of transcription via RNA polymerase II. RNA polymerase II catalyses the transcription of DNA to synthesise precursors of mRNA, snRNA and microRNA.<sup>57,58</sup> Numerous studies indicate an altered mRNA expression of proteins such as the serotonin 5-HT<sub>1A</sub> receptor,<sup>59,60</sup> brain-derived neurotrophic factor<sup>61–64</sup> and corticotrophin-releasing factor<sup>65,66</sup> in subjects suffering from MDD compared with controls. Furthermore, the evidence supporting a role for the dysfunction of microRNA-mediated regulated gene expression in MDD is increasing.<sup>67–71</sup> Therefore, CAG repeat variations in either *ATXN7* or *TBP* could modulate the function of RNA polymerase II, and thereby lead to changes in mRNA and microRNA expression that have previously been associated with MDD.<sup>59–71</sup> Hence, we intend to investigate the effect of CAG repeat length variations within the normal range in *ATXN7* and *TBP* on the expression and function of proteins thought to be associated with depression in a cellular model.

The presence of one subject with a CAG repeat size in the pathological range of *CACNA1A*, one subject with a CAG repeat size in the pathological range of *TBP* and two subjects with a pathological CAG repeat size in *ATXN2* within our two cohorts is remarkable. For these individuals, depression is apparently the first symptom with which SCA6, SCA17 or SCA2 manifests itself. The prevalence of an autosomal-dominant cerebral ataxia among Europeans is estimated to be between 1 and 3 per 100 000, of which 2% are diagnosed with SCA6,  $<1\%$  with SCA17 and 10% with SCA2,<sup>72</sup> resulting in a prevalence of 0.02–0.06 per 100 000,  $<0.01$ –0.03 per 10 000 and 0.10–0.30 per 100 000 for SCA6, SCA17 and SCA2, respectively. These numbers are in stark contrast to the substantially higher prevalence estimates in patients with depression in our study, 90 per 100 000 for SCA6, 50 per 100 000 for SCA17 and 50 per 100 000 for SCA2. Previously, we also found two individuals with incompletely penetrant *HTT* alleles (that is, those containing 36–39 CAG repeats) in this same population, both of whom suffered from depression.<sup>15</sup> The difference in proportion of individuals with intermediate or pathological CAG repeat lengths in PDAGs indeed tended to be higher in depressed compared with control subjects, although likely due to the relative

rarity of these expanded alleles the results did not reach statistical significance. Our findings are in line with another study that estimated the prevalence of incompletely penetrant *HTT* alleles in MDD patients to be ~3 in 1000, whereas such alleles were absent in the control group.<sup>73</sup> Together, these findings suggest that depression could be the first manifestation of polyglutamine diseases and that, conversely, polyglutamine diseases might be underdiagnosed in patients with depression.

In this study, we analysed a homogenous population by using samples from two Dutch cohorts for our analysis, thereby minimising the impact of population stratification. However, the use of a uniform group of individuals could also be seen as a limitation of our study, as it might have consequences for the generalisability of our findings. However, our results appear robust and consistent, although confirmation of our findings in other populations, as well as the elucidation of their pathophysiological basis, is warranted.

In conclusion, we observed a significant association between the risk of lifetime depression and CAG repeat size in *ATXN7* and *TBP*. A relatively large CAG repeat number in both alleles of either *ATXN7* or *TBP* substantially increased depression risk. Our findings add more critical evidence to the notion that repeat polymorphisms could act as complex genetic modifiers of depression and, therefore, could partially account for its 'missing heritability'. In addition, our findings indicate that the role of DNA repeat polymorphisms as potential genetic modifiers of other psychiatric disorders also needs further scrutiny.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

This study was supported by a VENI-grant (#91615080) from the Netherlands Organization of Scientific Research and a Marie Skłodowska-Curie Individual Fellowship grant from the European Union (Horizon 2020, #701130; NAA). The infrastructure for the NESDA study ([www.nesda.nl](http://www.nesda.nl)) has been funded through the Geestkracht programme of the Netherlands Organisation for Health Research and Development (Zon-Mw, grant number 10-000-1002) and participating universities (VU University Medical Center, Leiden University Medical Centre, University Medical Centre Groningen). The infrastructure for the NESDO study (<http://nesdo.amstad.nl>) is funded through the Fonds NutsOhra (project 0701-065), Stichting tot Steun VCVGZ, NARSAD The Brain and Behaviour Research Fund (grant ID 41080), and the participating universities and mental health-care organisations (VU University Medical Centre, Leiden University Medical Center, University Medical Centre Groningen, UMC St Radboud, and GGZ inGeest, GG Net, GGZ Nijmegen and Parnassia).

## REFERENCES

- Bromet E, Andrade LH, Hwang I, Sampson NA, Alonso J, de Girolamo G et al. Cross-national epidemiology of DSM-IV major depressive episode. *BMC Med* 2011; **9**: 90.
- Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* 2013; **382**: 1575–1586.
- Greenberg PE, Fournier AA, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry* 2015; **76**: 155–162.
- Major Depressive Disorder Working Group of the Psychiatric GC, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2013; **18**: 497–511.
- Okbay A, Baselmans BM, De Neve JE, Turley P, Nivard MG, Fontana MA et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat Genet* 2016; **48**: 624–633.
- Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR et al. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet* 2016; **48**: 1031–1036.

- Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med* 2010; **363**: 166–176.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J et al. Initial sequencing and analysis of the human genome. *Nature* 2001; **409**: 860–921.
- Hannan AJ. Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for 'missing heritability'. *Trends Genet* 2010; **26**: 59–65.
- Hannan AJ. TRPing up the genome: tandem repeat polymorphisms as dynamic sources of genetic variability in health and disease. *Discov Med* 2010; **10**: 314–321.
- Mirkin SM. Expandable DNA repeats and human disease. *Nature* 2007; **447**: 932–940.
- Shao J, Diamond MI. Polyglutamine diseases: emerging concepts in pathogenesis and therapy. *Hum Mol Genet* 2007; **16 Spec No. 2**: R115–R123.
- Bates G, Harper P, Jones L. *Huntington's Disease*, 3rd edn. Oxford University Press: New York, NY, USA, 2002.
- THSDCR G. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* 1993; **72**: 971–983.
- Gardiner SL, van Belzen MJ, Boogaard MW, van Roon-Mom WMC, Rosing MP, Albert M et al. Huntingtin gene repeat size variations affect risk of lifetime depression (submitted for publication).
- Orr HT, Chung MY, Banfi S, Kwiakowski TJ Jr., Servadio A, Beaudet AL et al. Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat Genet* 1993; **4**: 221–226.
- Chung MY, Ranum LP, Duvick LA, Servadio A, Zoghbi HY, Orr HT. Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nat Genet* 1993; **5**: 254–258.
- Matilla T, Volpini V, Genis D, Rosell J, Corral J, Davalos A et al. Presymptomatic analysis of spinocerebellar ataxia type 1 (SCA1) via the expansion of the SCA1 CAG-repeat in a large pedigree displaying anticipation and parental male bias. *Hum Mol Genet* 1993; **2**: 2123–2128.
- Pulst SM, Nechiporuk A, Nechiporuk T, Gispert S, Chen XN, Lopes-Cendes I et al. Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat Genet* 1996; **14**: 269–276.
- Sanpei K, Takano H, Igarashi S, Sato T, Oyake M, Sasaki H et al. Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nat Genet* 1996; **14**: 277–284.
- Giunti P, Sweeney MG, Harding AE. Detection of the Machado-Joseph disease/spinocerebellar ataxia three trinucleotide repeat expansion in families with autosomal dominant motor disorders, including the Drew family of Walworth. *Brain* 1995; **118**(Pt 5): 1077–1085.
- Maciel P, Gaspar C, DeStefano AL, Silveira I, Coutinho P, Radvany J et al. Correlation between CAG repeat length and clinical features in Machado-Joseph disease. *Am J Hum Genet* 1995; **57**: 54–61.
- Maruyama H, Nakamura S, Matsuyama Z, Sakai T, Doyu M, Sobue G et al. Molecular features of the CAG repeats and clinical manifestation of Machado-Joseph disease. *Hum Mol Genet* 1995; **4**: 807–812.
- Zhuchenko O, Bailey J, Bonnen P, Ashizawa T, Stockton DW, Amos C et al. Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. *Nat Genet* 1997; **15**: 62–69.
- Matsuyama Z, Kawakami H, Maruyama H, Izumi Y, Komure O, Uda F et al. Molecular features of the CAG repeats of spinocerebellar ataxia 6 (SCA6). *Hum Mol Genet* 1997; **6**: 1283–1287.
- David G, Abbas N, Stevanin G, Durr A, Yvert G, Cancel G et al. Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat Genet* 1997; **17**: 65–70.
- Benton CS, de Silva R, Rutledge SL, Bohlega S, Ashizawa T, Zoghbi HY. Molecular and clinical studies in SCA-7 define a broad clinical spectrum and the infantile phenotype. *Neurology* 1998; **51**: 1081–1086.
- Del-Favero J, Krols L, Michalik A, Theuns J, Lofgren A, Goossens D et al. Molecular genetic analysis of autosomal dominant cerebellar ataxia with retinal degeneration (ADCA type II) caused by CAG triplet repeat expansion. *Hum Mol Genet* 1998; **7**: 177–186.
- Johansson J, Forsgren L, Sandgren O, Brice A, Holmgren G, Holmberg M. Expanded CAG repeats in Swedish spinocerebellar ataxia type 7 (SCA7) patients: effect of CAG repeat length on the clinical manifestation. *Hum Mol Genet* 1998; **7**: 171–176.
- Nakamura K, Jeong SY, Uchihara T, Anno M, Nagashima K, Nagashima T et al. SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum Mol Genet* 2001; **10**: 1441–1448.
- Koide R, Kobayashi S, Shimohata T, Ikeuchi T, Maruyama M, Saito M et al. A neurological disease caused by an expanded CAG trinucleotide repeat in the TATA-binding protein gene: a new polyglutamine disease? *Hum Mol Genet* 1999; **8**: 2047–2053.
- Nagafuchi S, Yanagisawa H, Sato K, Shirayama T, Ohsaki E, Bundo M et al. Dentatorubral and pallidolysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. *Nat Genet* 1994; **6**: 14–18.



- 33 La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 1991; **352**: 77–79.
- 34 Bettencourt C, Hensman-Moss D, Flower M, Wiethoff S, Brice A, Goizet C *et al*. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann Neurol* 2016; **79**: 983–990.
- 35 Fan HC, Ho LI, Chi CS, Chen SJ, Peng GS, Chan TM *et al*. Polyglutamine (PolyQ) diseases: genetics to treatments. *Cell Transplant* 2014; **23**: 441–458.
- 36 Penninx BW, Beekman AT, Smit JH, Zitman FG, Nolen WA, Spinhoven P *et al*. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res* 2008; **17**: 121–140.
- 37 Wittchen HU, Robins LN, Cottler LB, Sartorius N, Burke JD, Regier D. Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). The Multicentre WHO/ADAMHA Field Trials. *Br J Psychiatry* 1991; **159**: 645–653, 658.
- 38 Comijs HC, van Marwijk HW, van der Mast RC, Naarding P, Oude Voshaar RC, Beekman AT *et al*. The Netherlands study of depression in older persons (NESDO); a prospective cohort study. *BMC Res Notes* 2011; **4**: 524.
- 39 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995; **57**: 289–300.
- 40 Lee SH, Goddard ME, Wray NR, Visscher PM. A better coefficient of determination for genetic profile analysis. *Genet Epidemiol* 2012; **36**: 214–224.
- 41 Andrade L, Caraveo-Anduaga JJ, Berglund P, Bijl RV, De Graaf R, Vollebergh W *et al*. The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *Int J Method Psychiatr Res* 2003; **12**: 3–21.
- 42 Okbay A, Baselmans BM, De Neve JE, Turley P, Nivard MG. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat Genet* 2016; **48**: 624–633.
- 43 Kashi Y, King D, Soller M. Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet* 1997; **13**: 74–78.
- 44 Kashi Y, King DG. Simple sequence repeats as advantageous mutators in evolution. *Trends Genet* 2006; **22**: 253–259.
- 45 Sawaya S, Bagshaw A, Buschiazio E, Kumar P, Chowdhury S, Black MA *et al*. Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. *PLoS ONE* 2013; **8**: e54710.
- 46 Sawaya SM, Lennon D, Buschiazio E, Gemmell N, Minin VN. Measuring microsatellite conservation in mammalian evolution with a phylogenetic birth-death model. *Genome Biol Evol* 2012; **4**: 636–647.
- 47 Lebre AS, Brice A. Spinocerebellar ataxia 7 (SCA7). *Cytogenet Genome Res* 2003; **100**: 154–163.
- 48 Pla P, Orvoen S, Saudou F, David DJ, Humbert S. Mood disorders in Huntington's disease: from behavior to cellular and molecular mechanisms. *Front Behav Neurosci* 2014; **8**: 135.
- 49 Chaudhuri KR, Schapira AH. Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment. *Lancet Neurol* 2009; **8**: 464–474.
- 50 Schmitz-Hubsch T, Coudert M, Tezenas du Montcel S, Giunti P, Labrum R, Durr A *et al*. Depression comorbidity in spinocerebellar ataxia. *Mov Disord* 2011; **26**: 870–876.
- 51 Rub U, Schols L, Paulson H, Auburger G, Kermer P, Jen JC *et al*. Clinical features, neurogenetics and neuropathology of the polyglutamine spinocerebellar ataxias type 1, 2, 3, 6 and 7. *Progr Neurobiol* 2013; **104**: 38–66.
- 52 Stevanin G, Brice A. Spinocerebellar ataxia 17 (SCA17) and Huntington's disease-like 4 (HDL4). *Cerebellum* 2008; **7**: 170–178.
- 53 Bauer P, Laccone F, Rolfs A, Wullner U, Bosch S, Peters H *et al*. Trinucleotide repeat expansion in SCA17/TBP in white patients with Huntington's disease-like phenotype. *J Med Genet* 2004; **41**: 230–232.
- 54 Helmlinger D, Hardy S, Sasorith S, Klein F, Robert F, Weber C *et al*. Ataxin-7 is a subunit of GCN5 histone acetyltransferase-containing complexes. *Hum Mol Genet* 2004; **13**: 1257–1265.
- 55 Wiczyk E, Brand M, Jacq X, Tora L. Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. *Nature* 1998; **393**: 187–191.
- 56 Lescure A, Lutz Y, Eberhard D, Jacq X, Krol A, Grummt I *et al*. The N-terminal domain of the human TATA-binding protein plays a role in transcription from TATA-containing RNA polymerase II and III promoters. *EMBO J* 1994; **13**: 1166–1175.
- 57 Kornberg RD. Eukaryotic transcriptional control. *Trends Cell Biol* 1999; **9**: M46–M49.
- 58 Sims RJ 3rd, Mandal SS, Reinberg D. Recent highlights of RNA-polymerase-II-mediated transcription. *Curr Opin Cell Biol* 2004; **16**: 263–271.
- 59 Lopez-Figueroa AL, Norton CS, Lopez-Figueroa MO, Armellini-Dodel D, Burke S, Akil H *et al*. Serotonin 5-HT1A, 5-HT1B, and 5-HT2A receptor mRNA expression in subjects with major depression, bipolar disorder, and schizophrenia. *Biol Psychiatry* 2004; **55**: 225–233.
- 60 Lopez JF, Chalmers DT, Little KY, Watson SJ. A.E. Bennett Research Award. Regulation of serotonin1A, glucocorticoid, and mineralocorticoid receptor in rat and human hippocampus: implications for the neurobiology of depression. *Biol Psychiatry* 1998; **43**: 547–573.
- 61 Cattaneo A, Gennarelli M, Uher R, Breen G, Farmer A, Aitchison KJ *et al*. Candidate genes expression profile associated with antidepressants response in the GENDEP study: differentiating between baseline 'predictors' and longitudinal 'targets'. *Neuropsychopharmacology* 2013; **38**: 377–385.
- 62 D'Addario C, Dell'Osso B, Galimberti D, Palazzo MC, Benatti B, Di Francesco A *et al*. Epigenetic modulation of BDNF gene in patients with major depressive disorder. *Biol Psychiatry* 2013; **73**: e6–e7.
- 63 Pandey GN, Dwivedi Y, Rizavi HS, Ren X, Zhang H, Pavuluri MN. Brain-derived neurotrophic factor gene and protein expression in pediatric and adult depressed subjects. *Progr Neuro-psychopharmacol Biol Psychiatry* 2010; **34**: 645–651.
- 64 Zhou L, Xiong J, Lim Y, Ruan Y, Huang C, Zhu Y *et al*. Upregulation of blood proBDNF and its receptors in major depression. *J Affect Disord* 2013; **150**: 776–784.
- 65 Austin MC, Janosky JE, Murphy HA. Increased corticotropin-releasing hormone immunoreactivity in monoamine-containing pontine nuclei of depressed suicide men. *Mol Psychiatry* 2003; **8**: 324–332.
- 66 Raadsheer FC, van Heerikhuizen JJ, Lucassen PJ, Hoogendijk WJ, Tilders FJ, Swaab DF. Corticotropin-releasing hormone mRNA levels in the paraventricular nucleus of patients with Alzheimer's disease and depression. *Am J Psychiatry* 1995; **152**: 1372–1376.
- 67 Geaghan M, Cairns MJ. MicroRNA and posttranscriptional dysregulation in psychiatry. *Biol Psychiatry* 2015; **78**: 231–239.
- 68 Hommers LG, Domschke K, Deckert J. Heterogeneity and individuality: microRNAs in mental disorders. *J Neural Trans* 2015; **122**: 79–97.
- 69 Hunsberger JG, Austin DR, Chen G, Manji HK. MicroRNAs in mental health: from biological underpinnings to potential therapies. *Neuromol Med* 2009; **11**: 173–182.
- 70 Issler O, Chen A. Determining the role of microRNAs in psychiatric disorders. *Nat Rev Neurosci* 2015; **16**: 201–212.
- 71 O'Connor RM, Grenham S, Dinan TG, Cryan JF. microRNAs as novel antidepressant targets: converging effects of ketamine and electroconvulsive shock therapy in the rat hippocampus. *Int J Neuropsychopharmacol* 2013; **16**: 1885–1892.
- 72 Durr A. Autosomal dominant cerebellar ataxias: polyglutamine expansions and beyond. *Lancet Neurol* 2010; **9**: 885–894.
- 73 Perlis RH, Smoller JW, Mysore J, Sun M, Gillis T, Purcell S *et al*. Prevalence of incompletely penetrant Huntington's disease alleles among individuals with major depressive disorder. *Am J Psychiatry* 2010; **167**: 574–579.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

© The Author(s) 2017

Supplementary Information accompanies the paper on the Translational Psychiatry website (<http://www.nature.com/tp>).